A SIMULATION OF RANDOM WALKS FOR USE AS AN EDUCATIONAL DEVICE

Thomas Richard Himstreet

Library
Naval Postgraduate School
Monterey, California 93940

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

A SIMULATION OF RANDOM WALKS
FOR USE AS AN EDUCATIONAL DEVICE

by

Thomas Richard Himstreet

Thesis Advisor: Richard W. Butterworth

March 1973

T153724

Approved for public release; distribution unlimited.

A Simulation of Random Walks
For Use as an Educational Device

by

Thomas Richard Himstreet
Lieutenant Commander, United States Naval Reserve
B. A. Simpson College, 1961

submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

Naval Postgraduate School
March 1973

Library
Naval Postgraduate School
Monterey, California 93940

ABSTRACT

The usefulness of random walks in mathematical modeling is often overshadowed by the problems that confront both students and instructors of probability. The counter-intuitive conclusions which arise produce both ambiguities and misunderstandings.

In this thesis, the techniques of computer simulation have been combined with the visual appeal of interactive graphic displays to develop a simulation of random walks. This simulation features interactive routines which are easy to use and take advantage of the insight and visual capabilities of the user to build an intuitive background of the subject matter. Statistics whose empirical distributions are asymptotically Arc Sine and Normal plus the gambler's ruin problem are displayed under various experimental conditions which the user designs.

This simulation is appropriate for use both by instructors to complement their classroom presentations and by students to enhance their understanding of the theory.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

The author is deeply indebted to a number of people for their special assistance. Professor R. W. Butterworth of the Naval Postgraduate School has been responsible for the development of several computer simulations of stochastic processes for educational purposes. Previous simulations have been done on Markov chains, renewal processes, and poisson processes. It was his suggestion that Random Walks might be suitable for computer simulation and graphic display. Hopefully, this simulation will be of use both to the students and instructors of probability at the Naval Postgraduate School.

A special note of thanks to Mr. R. Limes, Mr. W. Thomas, and Mr. A. Wong of the Naval Postgraduate School Electrical Engineering Computer Laboratory for the technical assistance they provided. Their contributions aided the author over innumerable pitfalls.

I. INTRODUCTION

Some of the simplest and most versatile of the stochastic processes are random walks. A random walk, as defined by Barr and Zehna (Ref. 1), is simply a point moving along the real number line in such a way that at each time $t (t = 0, 1, 2, \ldots)$, the point is an integer on the line. When the coordinates of the point are $(t, 0)$, a "equalization" or a "return to equilibrium" is said to have occurred.

Random walks are one-dimensional Markov chains and are used to model a wide range of phenomena, all characterized by random fluctuations. Such uses include exact models of simple coin-tossing games and first approximations to chance-dependent processes in Physics, Economics, and learning theory. Also included are the probabilistic development of Galton's rank order test, tests of the Kolmogorov-Smirnov type and sequential sampling.

Despite the simplicity of the structure of random walks and their wide range of uses, classroom presentations of the theory of random walks pose problems for both instructors and students of probability. These problems arise from the fact that although the basic premises of random walks can be developed quite simply from elementary combinatorial techniques, the conclusions encountered are often quite counter-intuitive in nature.

For example, if we visualize a random walk in terms of a simple coin-tossing game, the path followed by the random walk represents the cumulative number of heads minus tails at any point in time during the game. With probability one-half, no equalization will occur in the second half of the game regardless of the length of the game. In fact, it is quite likely that the winning player will be ahead for the entire

game. Also, the maximum accumulated gain, which is the highest value taken on by the cummulative number of heads minus tails during the game, is more likely to be attained at the very beginning or very end of the game.

The above aspects of random walks, exemplified by the simple coin-tossing game, are counter-intuitive to beginning probability students because of the preconceived notions concerning chance fluctuations and implications of the law of large numbers. Students often mistakingly assume that one coin tossed many times will yield the same statistical characteristics as an equal number of coins tossed once. The total of the heads minus tails is the same in both instances. It is the manner in which the single coin tossed many times arrives at this total that is the source of the student's misunderstanding.

A computerized simulation of random walks, combined with the capability to graphically display the results of the simulation, provides a very practical instructional device for dynamically representing some of the more subtle aspects of random walk theory. Interactive routines which feature graphic displays facilitate the development of student insight and comprehension by permitting the student to repeat specific simulations, vary parameters and to observe the effect on resulting displays. This helps to solidify and re-affirm his convictions on the subject matter. Instructors can use such a device equally well to complement their classroom presentations and to stimulate student interest in the topics by simulating and dynamically displaying examples of the material they are presenting.

II. DESCRIPTION OF A RANDOM WALK

A simple and intuitive characterization of a random walk is a coin-tossing game in which the primary statistic of interest is the cumulative number of heads minus tails. This total is denoted by x which represents the value of the cumulative number of heads minus tails at the time of the n^{th} toss of the coin. The event of a head occuring on a particular toss of the coin results in a unit increment to the cumulative number of heads minus tails. The probability of this event is called the success probability and is denoted by p. Decrements occur with probability $1-p$. Each change in the cumulative number of heads minus tails is the result of an independent trial with the success probability p.

If the success probability p is one-half, the random walk is said to be symmetric, otherwise it is said to be an unsymmetric random walk. Random variables will be displayed and discussed in terms of this characterization. Figure 1, on Page 9 , depicts a representative random walk. In this diagram, the cumulative number of heads minus tails is three after the fifth toss of the coin.

More formally stated, a random walk is defined as a polygonal line running from the origin to an arbitrary point (n,x) whose vertices have abscissas $0, 1, \ldots n$, and ordinates $S_0, S_1, \ldots S_n$ and where $S_0 = 0$, $S_k - S_{k-1} = \pm 1$ and $S_n = x$. For example, x represents the total number of $+1$'s minus the number of -1's. Referring to Figure 1, we see that at the point $(5,3)$, $S_5 = 3$.

In the symmetric case, there are 2^n equally likely paths of length n from any given point. From the origin to an arbitrary point (n,x) there are $N_{n,x} = \binom{n}{\frac{n+x}{2}}$ different paths, $N_{n,x} = 0$ when $\frac{n+x}{2}$ is not an

integer, $x < n$. The probability of visiting an arbitrary point x at time n denoted by $P_{n,x}$ is:

$$P_{n,x} = P\{S_n = x\} = N_{n,x} \cdot 2^{-n} = \binom{n}{\frac{n+x}{2}} 2^{-n} \qquad \text{(II.1)}$$

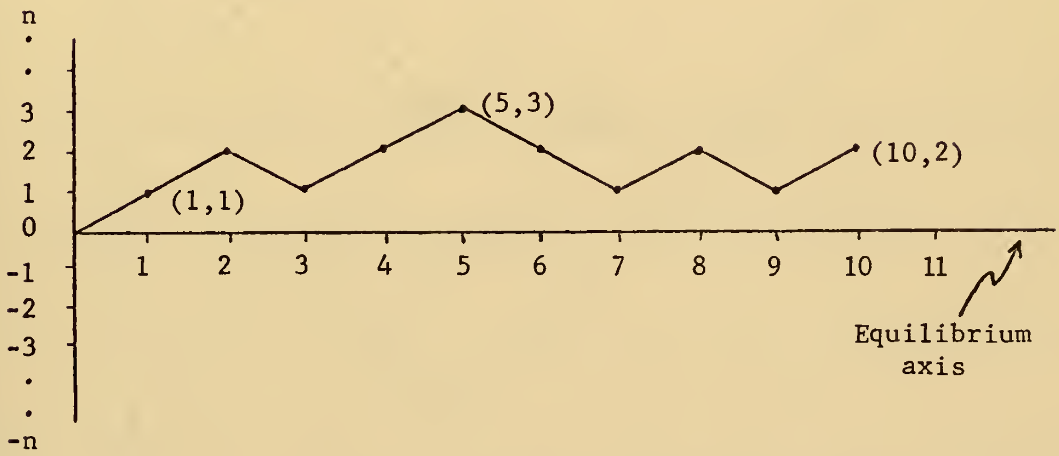

Figure 1. A Random Walk Illustrating the Ballot Theorem

The basic premises from which the theory of random walks is developed are the reflection principle, the ballot theorem, and the concept of a "return to equilibrium." The reflection principle involves a line known as the equilibrium axis. This line represents all points in the path of the random walk where the cumulative number of heads minus tails is equal to zero. The reflection principle states that the number of paths between any two points A and B lying above the equilibrium axis which either touch or cross the axis is equal to the total number of paths from A' to B, where A' is the reflection of A below the axis, for example, if A = (n,x), then A' = (n,-x). Figure 2 is a graphic representation of the reflection principle, with points C, D, and E lying on the equilibrium axis.

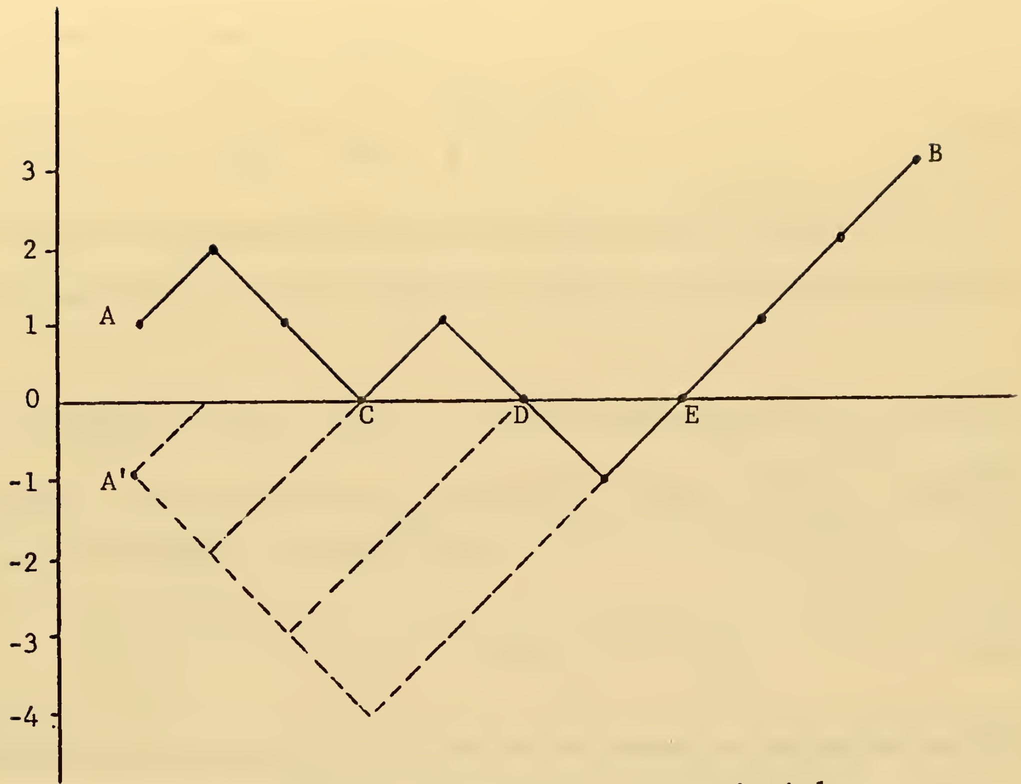

Figure 2. The Reflection Principle

The ballot theorem states that there are exactly $\frac{x}{n}N_{n,x}$ paths from the origin to the point (n,x) lying above the equilibrium axis which never touch or cross the axis. Specifically, those are the paths for which $S_0 = 0$, $S_1 > 0$, ... $S_n = x$. The random walk illustrated in Figure 1 is of this type with $n = 10$ and $x = 2$. The total number of such paths is the difference between the total number of paths from the point $(1,1)$ to the point (n,x), which is $N_{n-1,x-1}$, and the number of paths from $(1,1)$ to (n,x) which touch or cross the equilibrium axis, which is $N_{n-1,x+1}$ according to the reflection principle. The desired result is then $N_{n-1,x-1} - N_{n-1,x+1}$ which is equal to $\frac{x}{n} N_{n,x}$.

A "return to equilibrium", denoted by μ_k is said to have occurred at stage k if $S_k = 0$ where k is necessarily an even number. Letting $k = 2\nu$, the probability of this event occurring is given by formula II.1

with $n = \nu$ and $x = 0$ and is equal to

$$\mu_{2\nu} = P_{2\nu,0} = \binom{2\nu}{\nu} 2^{-2\nu} \tag{II.2}$$

In figure 2 on the preceeding page, Points C, D, and E represent "returns to equilibrium" where $k = 2\nu$ is equal to four, six and eight, respectively.

The probability that no return to equilibruim occurs up to and including stage 2ν is the same as the probability that a return to equilibrium occurs at stage 2ν. Symbolically,

$$P\{S_1 \neq 0, \ldots S_{2\nu} \neq 0\} = P\{S_{2\nu} = 0\} = \mu_{2\nu} \tag{II.3}$$

Paths which satisfy this criterion are either above or below the equilibrium axis from time 1 to 2ν, originating at the origin. There are $N_{2\nu-1,\, 2r-1} - N_{2\nu-1,\, 2r+1}$, $(r < \nu)$, paths which are positive and occur with probability $\frac{1}{2}P_{2\nu-1,1}$ and an equivalent number which are negative. Therefore:

$$\mu_{2\nu} = 2(\tfrac{1}{2})\, P_{2\nu-1,1}. \tag{II.4}$$

III. RANDOM VARIABLES OF INTEREST

A. SYMMETRIC RANDOM WALKS

In terms of a coin-tossing game, there are four random variables that illustrate symmetric random walk concepts and are suitable for simulation and graphic display. These are the cumulative number of heads minus tails, the last time the cumulative number of heads minus tails is equal to zero in the game, the number of tosses during which the cumulative number of heads minus tails is positive, and the toss at which the cumulative number of heads minus tails first reaches its maximum value in the game.

The cumulative number of heads minus tails is the amount by which one player is ahead and the other is behind in the game. Figure 3 illustrates this random variable. Illustrated is a game of ten coin tosses in which one player is ahead for the first seven tosses. The game is even after the eighth toss and the other player is ahead for the remaining two tosses.

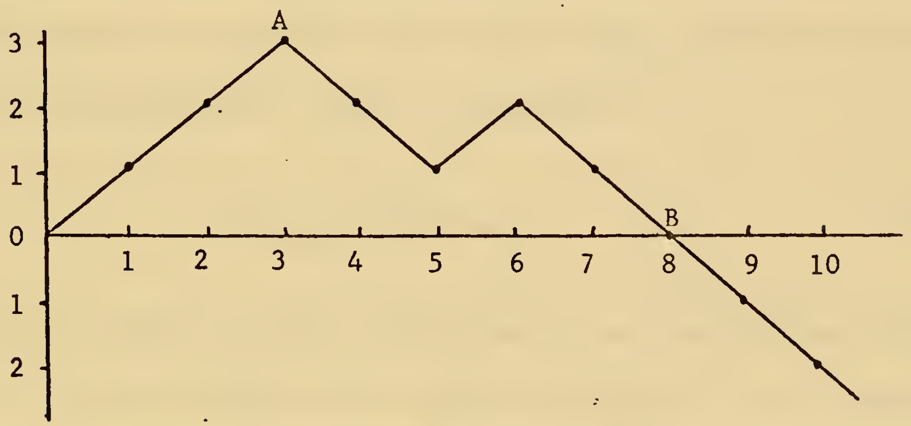

Figure 3. A Random Walk illustrating a first occurrance of the maximum (A) and a last return to equilibrium (B)

The cumulative number of heads minus tails (S_n) is the sum of n independent and identically distributed random variables, X_1, X_2, ..., X_n which can take the value of +1 or -1, each with probability $P = \frac{1}{2}$. The expected value of S_n is then $E(S_n) = (p-q)n = 0$ and the variance of S is $V(S_n) = 4pqn = n$.

The second random variable is the "last return to equilibrium" or the last time the total is equal to zero in the game. This occurs at the toss where the cumulative number of heads minus tails is equal to zero for the last time in the game. The Point B in Figure 3 on the preceeding page is illustrative of the last return to equilibrium in that particular random walk.

The probability that up to and including time 2n, the last return to equilibrium occurred at time 2k is denoted by $\alpha_{2k,2n}$. This is equivalent to the probability that a return to equilibrium occurs at time 2k and there are no subsequent returns in the remaining 2n-2k time units of the trial. Formulas II.2 and II.4 show these probabilities to be equal to μ_{2k} and μ_{2n-2k} respectively. Since the events are mutually independent, the desired probability is the product of the two individual probabilities. Hence, the formula is

$$\alpha_{2k,2n} = \mu_{2k}\,\mu_{2n-2k}. \tag{III.A.1}$$

The third random variable is the time (i.e. toss) during which the cumulative number of heads minus tails is positive. It is represented in Figure 3, Page 12, by the eight units between the origin and the Point B, where a return to equilibrium occurs. In general, the probability that up to and including toss 2n of the coin this total is positive for exactly 2k tosses, is denoted $b_{2k,2n}$.

The ways in which the number of heads minus tails can remain positive occurs in two ways. First, the cumulative number of heads minus tails may be positive for the first 2r tosses of the coin ($r < k$) and positive for exactly 2k-2r of the remaining 2n-2r tosses. Figure 4 illustrates an occurrance of this type.

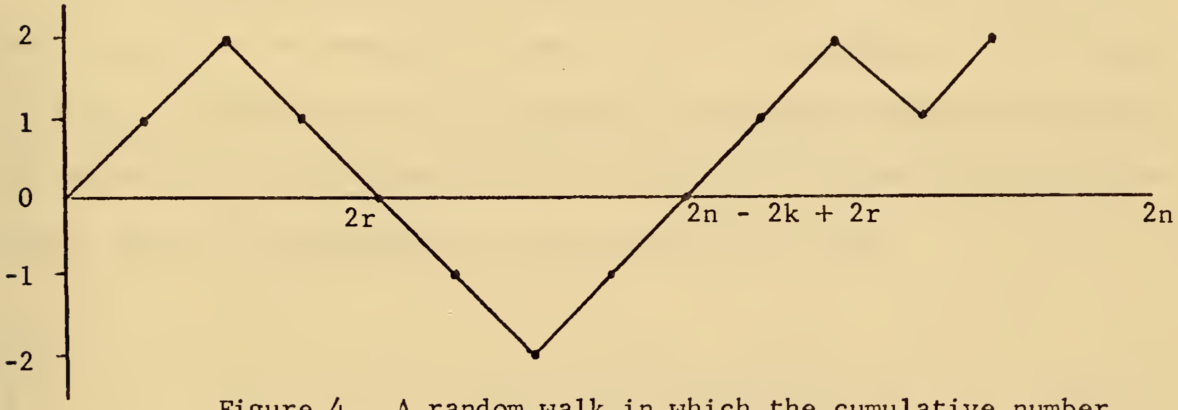

Figure 4. A random walk in which the cumulative number of heads minus tails is positive for the first 2r tosses and for exactly 2k-2r of the remaining 2n-2r tosses.

This event requires that the first return to equilibrium, denoted by f_{2r}, occurs on the $2r^{th}$ toss. The probability of a first return to equilibrium occurring at the $2r^{th}$ toss is equal to the probability that no return occurs in the first 2r-2 tosses, (μ_{2r-2}) minus the probability that a return to equilibrium occurs at the $2r^{th}$ toss (μ_{2r}). These follow from Formulas II.2, Page 11 and II.4, Page 11. As a result we have

$$f_{2r} = \mu_{2r-2} - \mu_{2r} \qquad \text{(III.A.2)}$$

In order for exactly 2k-2r of the remaining 2n-2r tosses to be positive, a return to equilibrium must occur at the $2n-2k^{th}$ toss and the remaining 2k-2r tosses must be positive. Formulas II.2 and II.4 again tell us that these probabilities are μ_{2n-2k} and $\frac{1}{2}\mu_{2k-2r}$

respectively. The total probability for this occurrence is the product of the three individual probabilities summed over all possible values of r.

$$\tfrac{1}{2}\mu_{2n-2k} \sum_{r=1}^{k} f_{2r}\, \mu_{2k-2r} = \tfrac{1}{2}\mu_{2k}\, \mu_{2n-2k} \qquad \text{(III.A.3)}$$

The second way in which the cumulative number of heads minus tails can be positive for exactly 2k tosses up to and including the $2n^{th}$ toss is for the first 2r tosses to be spent on the negative side of the axis and exactly 2k of the remaining 2n-2r tosses to be spent on the positive side. Figure 5 illustrates an occurrence of this type.

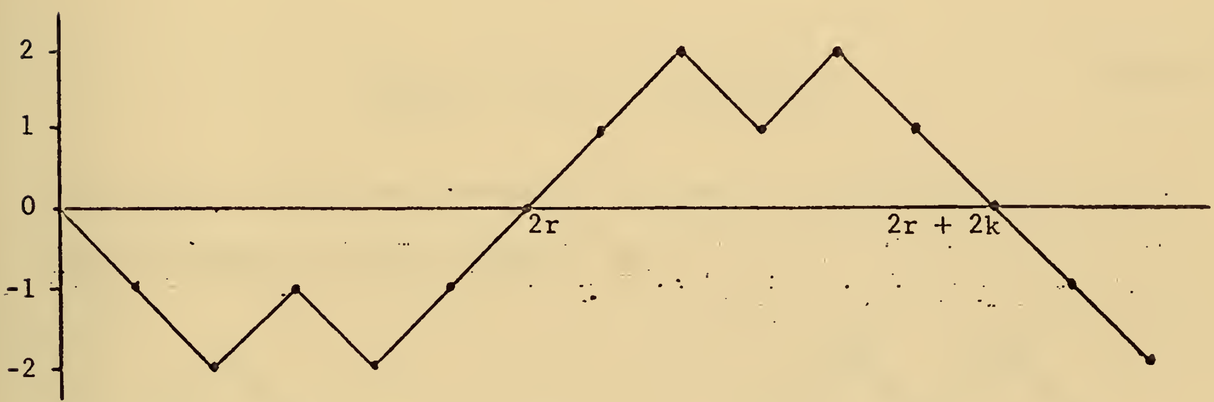

Figure 5. A Random Walk which is positive for 2k units.

This event requires a first return to equilibrium to occur at the $2r^{th}$ toss. This probability was just shown in Formula (III.2) to be $f_{2r} = \mu_{2r-2} - \mu_{2r}$. For the total to be positive for exactly 2k or the remaining 2n-2r tosses, a subsequent return to equilibrium must occur in the remaining 2n-2k-2r tosses. These probabilities, as given by Formulas (II.2) and (II.4), are μ_{2k} and $\mu_{2n-2k-2r}$ respectively. The total probability of this event then becomes

$$\tfrac{1}{2}\mu_{2k} \sum_{r=1}^{k} f_{2r}\, \mu_{2n-2k-2r} = \tfrac{1}{2}\mu_{2k}\, \mu_{2n-2k} \qquad \text{(III.A.4)}$$

The total probability that up to and including toss 2n of the game, the cumulative number of heads minus tails is positive for exactly 2k tosses of the coin is given by

$$b_{2k,2n} = \tfrac{1}{2}\mu_{2k}\,\mu_{2n-2k} + \tfrac{1}{2}\mu_{2k}\,\mu_{2n-2k} = \mu_{2n}\,\mu_{2n-2k}. \qquad \text{(III.A.5)}$$

The fourth random variable to be considered is the time (i.e. toss) where the cumulative number of heads minus tails first attains the maximum value for the game. This maximum is attained at the $2k^{th}$ toss if

$$S_0 < S_{2k},\ S_1 < S_{2k},\ \ldots\ S_{2k-1} < S_{2k} \qquad \text{(III.A.6)}$$

$$S_{2k+1} \le S_{2k},\ \ldots,\ S_{2n} \le S_{2k} \qquad \text{(III.A.7)}$$

In Figure 6, the first maximum is attained at Point A. This value is also attained at Points B and C.

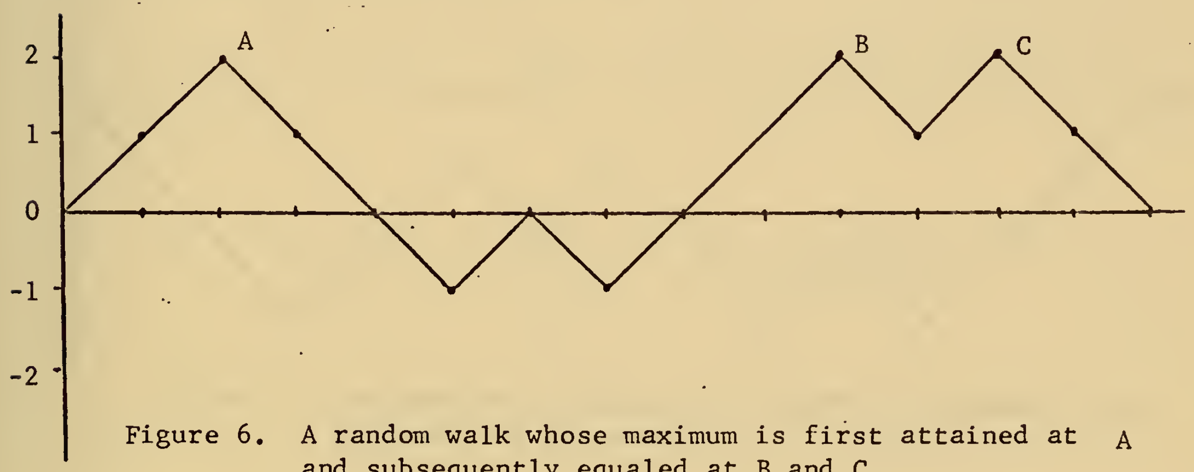

Figure 6. A random walk whose maximum is first attained at A and subsequently equaled at B and C.

For 0 < k < n, the event (III.6) occurs with probability $\tfrac{1}{2}\mu_{2k}$ and the event (III.7) occurs with probability μ_{2n-2k}. For k = 0 and k = n, these probabilities are μ_{2n} and $\tfrac{1}{2}\mu_{2n}$. Figures 7 and 8 illustrate these two concepts.

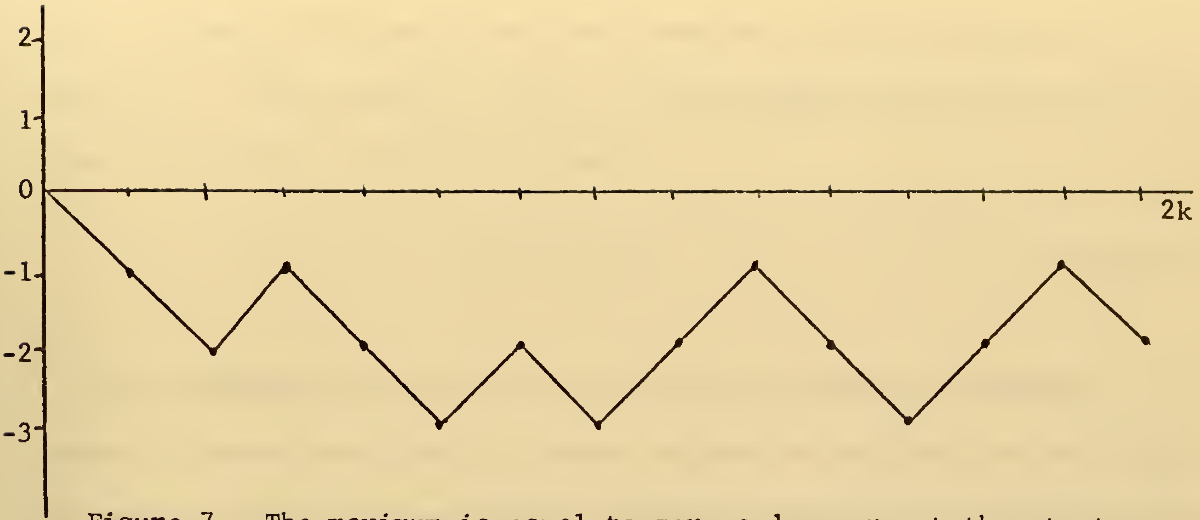

Figure 7. The maximum is equal to zero and occurs at the start of the game. i.e., k = 0

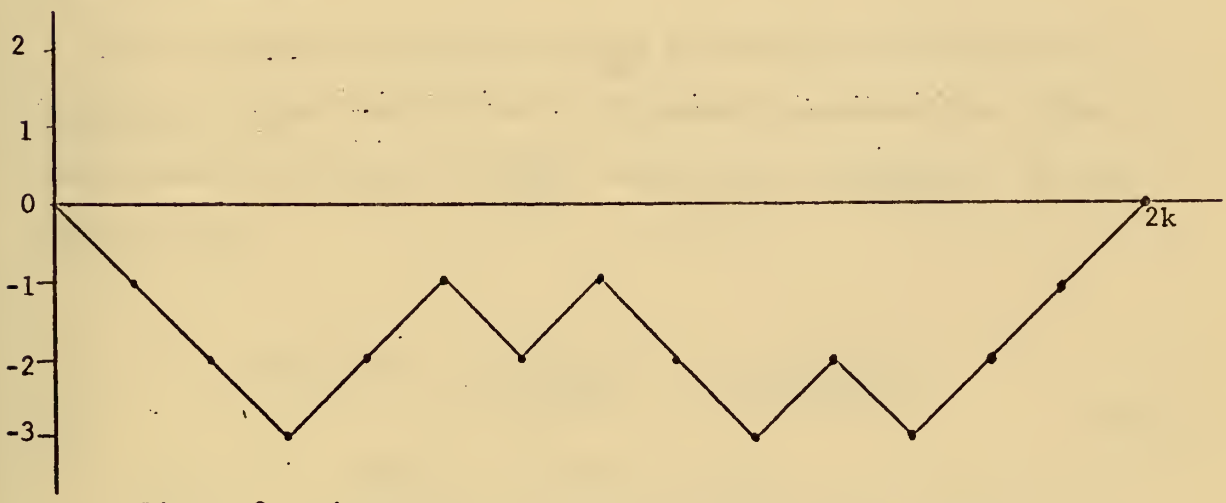

Figure 8. The maximum is equal to zero and is attained on the $2k^{th}$ toss of the coin.

These probabilities again come from Formulas (II.2) and (II.4) on Page 11. The probability that the first maximum occurs at the $2k^{th}$ toss of the coin is therefore equal to

$$\mu_{2k}\,\mu_{2n-2k}. \qquad \text{(III.A.8)}$$

The probability distributions of the last time the cumulative number of heads minus tails is equal to zero in the game, of the number of tosses on which the cumulative number of heads minus tails is positive, and of the toss on which the cumulative number of heads minus tails first takes on its maximum value are identical and all are equal to $\mu_{2k}\mu_{2n-2k}$. This distribution can be approximated by the Beta distribution with parameters $\alpha_1 = \alpha_2 = \frac{1}{2}$ which is more commonly referred to as the Arc Sine Distrbution.

The Arc Sine approximation to the probability distribution of $\mu_{2k}\mu_{2n-2k}$ evolves from the use of Stirling's approximation. This approximation gives $\mu_{2k} \approx \frac{1}{\sqrt{\pi k}}$ and $\mu_{2n-2k} \approx \frac{1}{\sqrt{\pi(n-k)}}$. It then follows that

$$\mu_{2k}\,\mu_{2n-2k} \approx \frac{1}{\pi\sqrt{k(n-k)}} = \frac{1}{n\pi\sqrt{x(1-x)}}$$

$$\text{where } x = \frac{k}{n} \text{ and } 0 < x < 1 \qquad \text{(III.A.9)}$$

which is the probability density function of the Arc Sine distribution. The mean of this distribution is $\frac{1}{2}$ and the variance is 1/8.

B. UNSYMMETRIC RANDOM WALKS

Unsymmetric random walks are most easily illustrated by drawing analogies to the well known gambler's ruin problem. In this context, two players, each with an initial level of capital and betting unit amounts on each play, compete in a game of chance until one of the player's resources are exhausted. Player one represents a gambler with limited initial capital resources, z. The second player is referred to as the "adversary." His initial resources, a-z, where a denotes the total amount of capital in the game, may be either finite or infinite. Each play is an independent trial with common probability of success p for player one.

Statistics of interest include the probability of player one being ruined, the expected duration of the game, and the expected gain of player one in the game. The probability of ruin for player one is denoted by q_z and is derived by conditioning on the result of the first play of the game. After the first play, the gambler's capital is z + 1 with probability p and z-1 with probability q. The resulting difference equation is

$$q_z = p\, q_{z+1} + q\, q_{z-1} \qquad 1 \le z \le a - 1 \qquad \text{(III.B.1)}$$

Boundary conditions are $q_0 = 1$ and $q_a = 0$. $q_0 = 1$ states that with probability one, player one is ruined when his capital is depleted. $q_a = 0$ states that the probability of player one being ruined when he owns all the capital is equal to zero. If $p \neq q$, the solution to the difference equation is

$$q_z = \frac{(q/p)^a - (q/p)^z}{(q/p)^a - 1} \qquad \text{(III.B.2)}$$

If $p = q$, the solution is

$$q_z = 1 - \frac{z}{a} \qquad \text{(III.B.3)}$$

This derivation follows that given by Feller in Reference 2.

Playing against an infinitely rich adversary implies that $a \to \infty$. In this case

$$q_z = \begin{cases} 1 & \text{if } p \leq q \\ (q/p)^z & \text{if } p > q \end{cases} \qquad \text{(III.B.4)}$$

The expected gain in terms of player one's winnings is then the total capital in the game times the probability of player one winning $(1-q_z)$ minus his initial capital. The expected gain is denoted by G.

$$E[G] = a(1 - q_z) - z \qquad \text{(III.B.5)}$$

The expected duration of the game is again derived by conditioning on the outcome of the first event. Let D_z denote the expected duration of a game which begins with player one having capital z. Player one, by winning the first play, extends the length of the game by one event and a new game with capital $z + 1$ begins. By losing the first play he extends the game by one event and a new game with capital $z - 1$ begins. This leads to the difference equation

$$D_z = p\, D_{z+1} + q\, D_{z-1} + 1 \qquad z = 1, 2, \ldots a - 1 \qquad \text{(III.B.6)}$$

For this equation, the boundary conditions are $D_0 = 0$ and $D_a = 0$. These conditions reflect the assumption that the game ends when the capital is

exausted. When $p \neq q$, the expected duration of the game is found to be

$$D_z = \frac{z}{q-p} - \frac{a}{q-p} \cdot \frac{1-(q/p)^z}{1-(q/p)^a} \tag{III.B.7}$$

For $p = q$, the expected duration of the game is

$$D_z = z(a - z). \tag{III.B.8}$$

IV. SIMULATION OF A RANDOM WALK

The program simulates a random walk and collects statistics on the four variables defined in Section III.A. The first three are the "last return to equilibrium", the number of time units in the trial during which the cumulative number of heads minus tails is positive, and the point in time during the trial where the maximum gain is first attained, given by Formulas III.A.1, 5, and 8 respectively. The fourth statistic collected is the cumulative number of heads minus tails which was shown to be asymptotically distributed Normal.

The program also contains a special segment that simulates unsymmetric random walks in terms of the gambler's ruin problem as described in Section III.B. Statistics are collected for player one's probability of ruin, expected gain in the game, and the expected duration of the game. Formulas III.B.2, 3, 4, 5, 7, and 8 represent the theoretical values of the statistics.

The program is modularly designed so that each statistic collected and displayed is a separate subroutine which operates independently from the remainder of the program. The only exception is the input-output functions which are carried out in a separate module common to all subroutines.

The user interacts with the program via the teletype-writer connected to the graphic display unit. Options are selected, parameters are specified, and the simulation is carried out entirely under the control of the user. The executing time of a particular run may be excessive due to the parameters which were specified for the simulation. When the possibility of this exists, the user is given the opportunity to

re-assign the parameters prior to executing the program. This enables the user to avoid execution of a simulation where the parameters have been erroneously specified.

Statistics collected whose empirical distributions are asympototically Arc Sine or Normal feature two separate displays. The random walk is dynamically displayed as it develops with a frequency plot displayed below. The frequency plot is simply an empirical density for the random variable being illustrated. This illustrates the development of the probability distribution as it evolves over many repetitions of the experiment. The number of events which make up each trial, the number of trials, and the number of data cells into which the statistics will be grouped are all user specified parameters. Figure 9 illustrates this display.

The second display repeats the frequency plot. The theoretical probability distribution is overlayed and a tabulation of sample and theoretical statistics for the expected value and standard deviation is made. Figures 10 and 11 illustrate the different displays of symmetric random walks.

The unsymmetric random walk case illustrated by the gambler's ruin problem displays the random walk as it evolves for each trial of the game. Additional axes are drawn representing the initial amount of capital possessed by player one and his adversary. The axis representing the adversary's initial capital is the barrier which, when reached by the random walk, is interpreted as a victory for player one. The equilibrium axis represents the point at which player one's capital is gone and is interpreted as a loss for player one. Theoretical and sample values are tabulated and displayed for the probabilities of

player one's ruin as defined by Formulas III.B.2, 3, and 4. His expected gain is shown by Formula III.B.5 and the expected duration of the game is defined by Formulas III.B.7 and 8. Figure 12 illustrates the display created by the gambler's ruin problem.

This program was written for execution on the XDS-9300 Computer and the Adage-10 Graphics Display Terminals, located at the Naval Postgraduate School in the Electrical Engineering Laboratory, Spangel Hall, Room 500. The programming language used was Scientific Data Systems (SDS) FORTRAN IV. Program execution is carried out under the standard XDS-9300 operating system and the Program Graphics and Text Editor (GATED). The program GATED is a specially developed software interface between the XDS-9300 Computer and the ADAGE-10 graphics display terminals and is not readily compatible with other operating systems. References 3 and 4 are technical manuals for SDS FORTRAN IV and GATED, respectively.

No prior knowledge of data processing is required to utilize this program. Operation of data processing equipment in the laboratory is the responsibility of the individual user, but step-by-step instructions are available for all equipment required. During normal working hours, laboratory technicians are available to assist the user.

Instructions for user interaction with the program are an additional option which can be accessed from the "menu page" in the same manner as any other option. These instructions are not meant to provide any theoretical information in support of the simulation presented, but serve to delineate acceptable input parameters and to clear up any ambiguities which may arise in the user's first experiences with the program.

Design features, such as those delineated by Reference 5, are incorporated into the program to make the system as user orientated as possible. All input parameters are validated, the user's needs are anticipated wherever they were deemed necessary. Reference 6 lists acceptable response times for various situations which arise in user-orientated interactive graphics systems.

Maximum benefits of this simulation can be attained if an instructor exposes his class of probability students to the simulation after first lecturing on the subject of random walks. As the class lecturers progress in the theory, students can experiment with the simulation under varying parameters on an individual or group basis. The program then complements the material presented in class and reinforces the students' knowledge on the subtleties of the theory.

The computer program for this simulation is completely annotated with comment cards to explain the program. Variables common to the entire program are defined in the first segment of the program which is the input-output segment. Variables common to individual subroutines are defined in their respective routines as appropriate.

Utilization of interactive computer graphics for instructional purposes is in its infancy at the Naval Postgraduate School. Professor R. Butterworth of the Operations Research and Administrative Sciences Department has supervised the development of programs which illustrate point processes, Markov chains, queuing models and this thesis on random walks. The use of interactive computer graphics is not restricted to illustrating probabilistic models alone, but might be useful in illustrating other Operations Research techniques. For example, reliability models, war gaming, and regression estimates could be areas for future research and development.

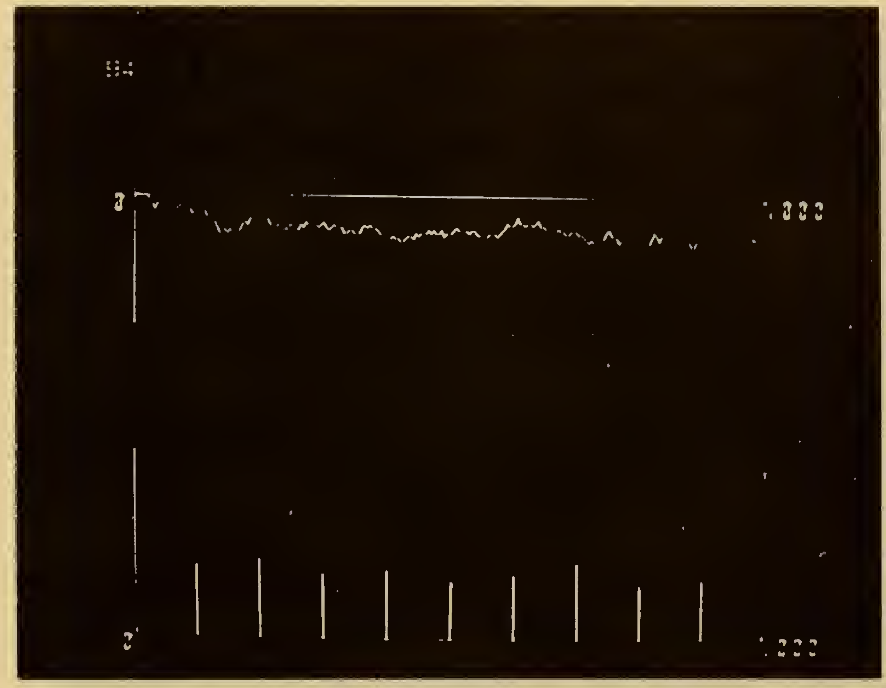

Figure 9. Graphic display of a random walk and frequency plot.

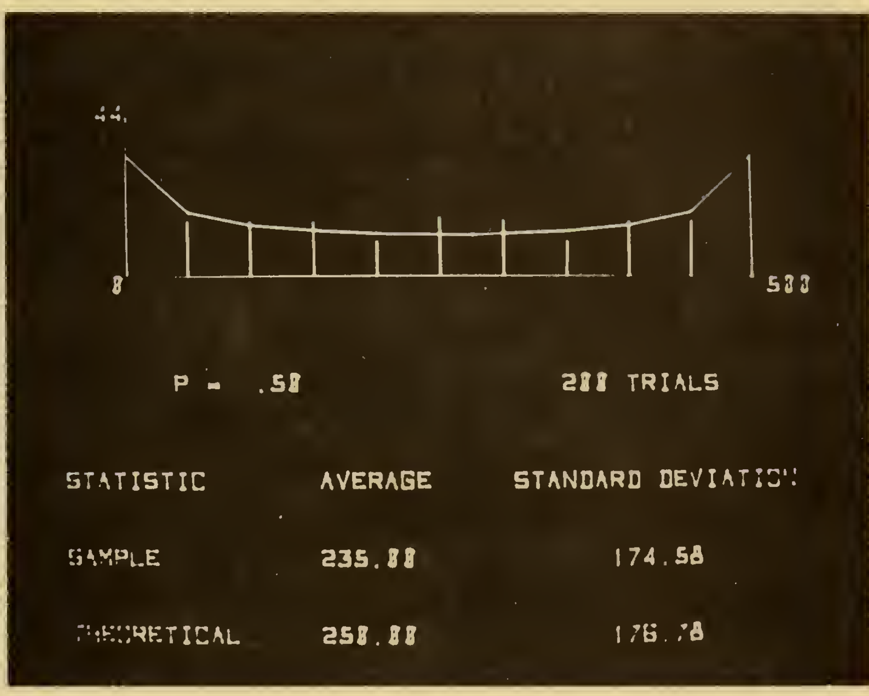

Figure 10. Frequency plot with Arc Sine distribution overlayed and statistics tabulated.

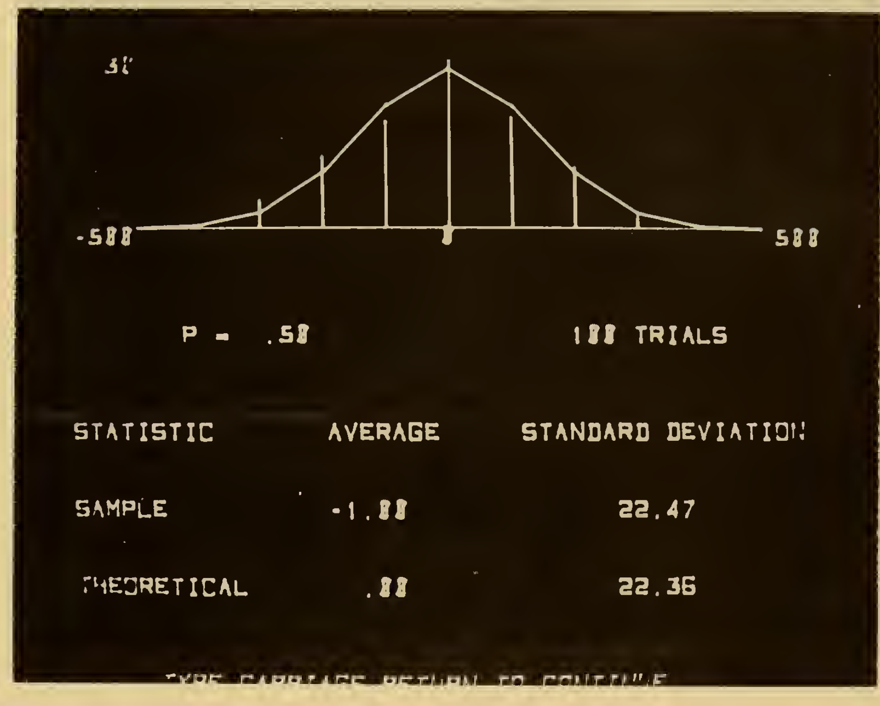

Figure 11. Frequency plot with normal distribution overlayed and statistics tabulated.

Figure 12. Graphic display of the gambler's ruin.

BIBLIOGRAPHY

[1] Barr, D. R. and Zehna, P. W., Probability, p. 335-336, Brooks/Cole, 1971.

[2] Feller, W., An Introduction to Probability Theory and Its Applications, V. 2, 3rd ed., p. 67-95, 342-355, Wiley, 1968.

[3] Scientific Data Systems Reference Manual, FORTRAN IV, October 1966.

[4] Naval Postgraduate School Electrical Engineering Computer Laboratory, Manual for the FORTRAN User, by R. D. DeLaura, 10 February 1968.

[5] Smith, L. B., "The Use of Interactive Graphics to Solve Numerical Problems," Communications of the ACM, V. 13, No. 10, p. 625-634, October 1970.

[6] Miller, R. B., "Response Time in Man-Computer Conversational Transactions," AFIPS FJCC, 1968, p. 267-277.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Chief of Naval Personnel Pers 11b Department of the Navy Washington, D. C. 20370	1
2.	Naval Postgraduate School, Code 55 Department of Operations Research and Administrative Sciences Monterey, California 93940	1
3.	Defense Documentation Center Cameron Station Arlington, Virginia 22314	2
4.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
5.	Professor R. W. Butterworth, Code 55BD Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
6.	Professor P. A. W. Lewis, Code 55LW Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
7.	Professor P. W. Zehna, Code 55ZE Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
8.	LCDR B. R. Webster, Code 55WT Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
10.	LCDR Thomas R. Himstreet 810 West Detroit Avenue Indianola, Iowa 50125	1

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

A Simulation of Random Walks For Use as an Educational Device

4. DESCRIPTIVE NOTES (Type of report and, inclusive dates)

Thesis, Master of Science, March 1973

5. AUTHOR(S) (First name, middle initial, last name)

Thomas R. Himstreet

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
March 1973	33	6

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO.	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Naval Postgraduate School Monterey, California 93940

13. ABSTRACT

The usefulness of random walks in mathematical modeling is often overshadowed by the problems that confront both students and instructors of probability. The counter-intuitive conclusions which arise produce both ambiguities and misunderstandings.

In this thesis, the techniques of computer simulation have been combined with the visual appeal of interactive graphic displays to develop a simulation of random walks. This simulation features interactive routines which are easy to use and take advantage of the insight and visual capabilities of the user to build an intuitive background of the subject matter. Statistics whose empirical distributions are asymptotically Arc Sine and Normal plus the gambler's ruin problem are displayed under various experimental conditions which the user disigns.

This simulation is appropriate for use both by instructors to complement their classroom presentations and by students to enhance their understanding of the theory.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Random Walk						
Simulation						
Educational Device						

2 JUN 76 24425

Thesis 143264
H536 Himstreet
c.1 A simulation of random walks for use as an educational device.

2 JUN 76 24425

Thesis 143264
H536 Himstreet
c.1 A simulation of random walks for use as an educational device.

thesH536
A simulation of random walks for use as
3 2768 001 01429 3
DUDLEY KNOX LIBRARY